WEST Search History

Hide Items | Restore | Clear | Cancel

DATE: Friday, January 21, 2005

Hide?	Set Name	Query	Hit Count
		DB=USPT; PLUR=NO; OP=ADJ	
☐	L20	l18 and l4 and (l3 or l2 or l12)	89
☐	L19	l10 and l18	14
☐	L18	L17 or l16 or l15 or l14	63325
☐	L17	345/$.ccls.	33372
☐	L16	715/$.ccls.	12391
☐	L15	707/$.ccls.	14193
☐	L14	705/$.ccls.	9749
☐	L13	L12 and l10	5
☐	L12	(categor$9).ab.	2877
☐	L11	L10 and (l1 or l3 or l4)	7
☐	L10	categoriz$9 same (page or HTML) same (goverment or medical or education or sports or history or entertainment)	23
☐	L9	l1 and l3 and l4	21
☐	L8	l3 same l4	2
☐	L7	L6 and (l3 or l4)	0
☐	L6	l1 and l2	3
☐	L5	l1 and l2 and l3 and l4	0
☐	L4	(assign$6 or tag or mark) near4 page$1	4692
☐	L3	list near3 categor$7	1497
☐	L2	(categorizing near4 (page or site))	16
☐	L1	(707/1 or 707/7 or 707/10 or 707/100).ccls.	6272

END OF SEARCH HISTORY

h e b b cg b chh e e g f ff e ch e g b

First Hit Fwd Refs Previous Doc Next Doc Go to Doc#

End of Result Set

Generate Collection Print

L6: Entry 3 of 3 File: USPT Nov 16, 1999

DOCUMENT-IDENTIFIER: US 5987457 A
TITLE: Query refinement method for searching documents

Current US Cross Reference Classification (1):
707/10

CLAIMS:

13. A method for refining an initial query phrase to search for web pages on the world wide web that are of interest to a user, comprising the steps of:

categorizing at least one web page found in a search using the initial query phrase as of interest based upon feedback from the user;

categorizing at least one other web page found in the search using the initial query phrase as not of interest based upon feedback from the user;

generating a list of keywords by analyzing only the categorized web pages;

ranking as first keywords, the keywords in the list of keywords which occur in only the web pages of interest;

ranking as second keywords, the keywords in the list of keywords which occur in only the web pages not of interest;

forming a refined query phrase to search for web pages which include one or more of a plurality of the highest ranked first keywords, and to filter out web pages which include any one or more of a plurality of the highest ranked second keywords.

Previous Doc Next Doc Go to Doc#

L8: Entry 1 of 2 File: USPT Jul 13, 2004

DOCUMENT-IDENTIFIER: US 6763496 B1
TITLE: Method for promoting contextual information to display pages containing hyperlinks

Detailed Description Text (57):
The category list components are used to automatically generate a list of one or more hyperlinks to documents on a web that are assigned a category matching the category associated with each category list component. For instance, suppose that a user has created three pages corresponding to the "large" category, including: elephant.htm, rhino.htm, and hippo.htm, and three pages corresponding to the "cats" category, including: lion.htm, tiger.htm, and leopard.htm. Each of these pages has an associated contextual information file containing meta-data entries, as shown in FIG. 9C. These contextual information files include an elephant.htm file 558, rhino.htm file 560, hippo.htm file 562, lion.htm file 564, tiger.htm file 566, and leopard.htm file 568. Each of these contextual information files contains a category meta-data entry that is used to assign a category to the page (the HTML document) with which the contextual information file is associated. For example, the "large" category is assigned to the HTML documents (not shown) that are associated with contextual information files 558, 560, and 562, and the "cats" category is assigned to the HTML documents (not shown) that are associated with contextual information files 564, 566, and 568. The category meta-data entries are preferably added to a contextual information file when its associated document is saved, as described above. The categories can be explicitly defined by the user, or implicit as part of some other process (such as a pre-save scan of the document for keywords). A given document may be assigned to one or more categories, or none at all.

Detailed Description Text (58):
When a design page is saved, an HTML document is created (or modified) that contains the HTML code (and JAVA script, as applicable) for displaying the design page on a browser. At this point, the data promotion engine is invoked to generate hyperlinks that correspond to each of the category list components in a given design page. The data promotion engine parses through the content of the design page document in search of category_bot entries. When the data promotion engine comes to a "category_bot" entry, it parses through the contextual information files on the site to identify any documents that are assigned to a category matching the category indicated by the category_bot entry. The data promotion engine then generates the HTML code to insert hyperlinks into the pages that have been assigned to the matching category.

Detailed Description Text (60):
The hyperlinks that are created on the display page (corresponding to the design page) are positioned relative to the location of the category list components on the design page. For example, FIG. 9B shows a display page 584, which corresponds to design page 550, as viewed on a browser 586. Hyperlinks 588 correspond to pages that have been assigned to the "large" category, while the hyperlinks 590 correspond to pages that have been assigned to the "cats" category.

Detailed Description Text (61):
Another feature of the category association scheme is the ability to automatically

h e b b g e e e f c e e g e ge

promote new hyperlinks to design pages when new pages are created and (or existing pages are) assigned to categories that correspond to category list components in the design page, without requiring the design page to be edited by a user so as to include the new hyperlinks. When a new page is created and saved, its author has the option of assigning a category to it. Alternately, an author can assign a category to an existing page or modify the category already assigned to an existing page. If a category is assigned to the new or existing page, the category information is stored as a meta-data entry in the contextual information file associated with the new or existing page, and the data promotion engine then parses through all of the documents on the site in search of documents that contain a category list component matching the category of the new document. The data promotion engine opens the matching files and adds HTML code to these files to add a hyperlink to the new or existing document.

Previous Doc Next Doc Go to Doc#

h e b b g e e e f c e e g e ge

L11: Entry 6 of 7 File: USPT Aug 12, 2003

DOCUMENT-IDENTIFIER: US 6606659 B1
TITLE: System and method for controlling access to internet sites

Detailed Description Text (7):
Embodiments of the system also provide methods for automatically categorizing Internet pages to create and update a database of categorized sites. This categorized database is then used within an Internet access control system to control user's access to Internet sites within certain categories. For example, if the system described herein assigns a particular Internet page to a "Sports" category, users that are restricted from viewing sports pages on the Internet will not be granted access to the requested site. In one embodiment, the system is installed within an Internet Gateway computer that controls traffic from the user to the Internet. Because the system described herein becomes more accurate with each page that is scored, minimal user intervention is required to assign pages to categories.

Detailed Description Text (34):
As discussed below, the determination of whether to assign a retrieved page to a particular category is made by comparing the page's relevance score for a particular category with a predetermined alpha value. If the page relevance score is higher than the alpha value for the category, the page is assigned to that category. If the score is lower than the alpha value, but greater than a beta value, the page is forwarded to a manual scoring system wherein technicians view the retrieved page and determine whether or not to include the page within the category. If the relevance of the page for a category is below the beta value, the page address is stored to a database of analyzed sites, and the system continues to score additional addresses.

Detailed Description Text (37):
In addition to the word identification table 200 is a category identification table 205 that provides a category ID number for each category within the system. The category identification table 205 also includes an alpha and beta score that provide the cut-off values for assigning a particular page to the selected category. For example, as illustrated in FIG. 3, the Sports category includes an alpha score of 920 and beta score of 810. If an Internet page is found to have a page relevance score of greater than 920 for the Sports category, it will be assigned to the Sports category. However, if the Internet page is found to have a page relevance score of between 810 and 920, it will be flagged for manual follow-up by a technician to determine whether or not it belongs within the Sports category. If the Internet page is found to have a page relevance score of below 810 for the Sports category, then it will not be flagged as being related to the Sports category. By using these values, the system determines whether or not to assign a particular page to one of the predefined categories.

Detailed Description Text (50):
However, if an address match between the requested address and the categorized database is found, the process 300 moves to a decision state 315 wherein a determination is made whether the current user has restricted access rights to specific categories of Internet pages. This determination can be made by reference to a list of network users, and an associated permissions table for each category

found within the categorized database. Thus, a particular user may be restricted from access to all Sports and Pornography categories but not restricted from Internet Commerce or Travel categories. An exemplary list of Internet categories is provided below in Table 1.

Detailed Description Text (70):
Referring to FIG. 7, a process 500 for creating the word relevance table 210 within the training database 125 is described. The process 500 begins at a start state 502 and then moves to a state 504 wherein a first category to train is selected. The category might be, for example, the Sports category. The process 500 then moves to a state 508 wherein web pages that have been predetermined to be within the chosen category (e.g., sports) are retrieved. Thus, because these pages are known to be within the category selected at state 504, the relevance of each word pair and word adjacency within the chosen page can be assigned a high relevance to the current category.

Detailed Description Text (73):
The process then moves to a state 530 wherein the current score for each word pair and word adjacency (1000) is averaged with the same word pair and word adjacency scores already stored in the word relevance table. Thus, if we are training the Sports category, and the word adjacency "Cleveland Browns" is found within the current page, it might be assigned a word adjacency value of 105 in the Sports category. However, if the term "Cleveland Browns" is already scored within the Sports category at a value of 89, the 105 value and the 85 value would be averaged to normalize the word adjacency score to the Sports category. This system therefore allows words that are used over and over within certain categories to be "up-trained" so that their relevance score with the chosen category will go up as they appear on more pages that are scored. In addition, it should be understood that the system is capable of parallel processing of a plurality of sites simultaneously.

Detailed Description Text (76):
Through the process 500 described above, a word relevance table is developed which includes normalized word relevances for every word pair and word adjacency that might be found in an Internet page. By analyzing new pages and by adding together the relevances of each word within the page, an automated system is provided for assigning a page relevance score for a particular page to each of the predetermined categories within the system. Thus, once a particular category has been trained by analysis of a large number of pages, the system can rapidly analyze new pages for their relevance to each of the predetermined categories. As described above in FIG. 2, a page retrieval module 110 is utilized for retrieving new Internet pages and sending them to the analysis module 120 for scoring.

Detailed Description Text (90):
Referring now to FIG. 11, a timer quota process 850 is illustrated. The timer quota process 850 begins at a start state 852 and then moves to a state 854 wherein a request is received for an Internet page or site. A determination of the category of the page or site is then made at a state 858 by reference to the categorized database 30. The process 850 then moves to a state 860 wherein any timer quota parameters for the selected category of sites are retrieved. For example, a quota parameter indicating that users can only spend, for example, 30 minutes within the Sports category might be retrieved at the state 860.

Previous Doc Next Doc Go to Doc#